molecular systems biology

Alternative TSSs are co-regulated in single cells in the mouse brain

Kasper Karlsson, Peter Lönnerberg, Sten Linnarsson

Corresponding author: Sten Linnarsson, Karolinska Institutet

Review timeline:

Submission date:	07 October 2016
Editorial Decision:	02 December 2016
Revision received:	02 March 2017
Editorial Decision:	03 April 2017
Accepted:	11 April 2017

Editor: Maria Polychronidou

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 02 December 2016

Thank you again for submitting your work to Molecular Systems Biology. We have heard back from two of the three referees who agreed to evaluate your manuscript. As you will see below, the reviewers raise substantial concerns on your work, which unfortunately preclude the publication of the study in its current form.

In particular, the reviewers mention that additional analyses are required i) to better support the main findings and ii) to provide conclusive novel insights. However, considering that they appreciate that the addressed topic is timely and important for the field, we would like to offer you a chance to revise and extend the study and address the points raised.

Without repeating all the points listed below, the most fundamental issues that need to be addressed are the following:

- Additional analyses should be performed to better support the existence of multiple TSSs. Both reviewers provide constructive suggestions in this regard.

- Follow up analyses on the reported findings (i.e. Cst3 alternative TSS preference in the different cell types) would significantly enhance the impact of the study. Reviewer #2 provides suggestions and lists (see point #2) other findings that could alternatively be followed up in order to provide novel insights.

REFEREE REPORTS

Reviewer #1:

The manuscript entitled "Alternative TSSs are co-regulated in single cells in the mouse brain" is a follow up analysis of previously published Zeisel A., et al "Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq". Understanding of TSS usage in single cells and also across different cell types is essential step to dissect the regulatory process surrounding promoters and this manuscript uses the single cell sequencing data to address this important question. The authors report that alternative TSS usage is a regulated process and the correlation between two TSS expressed in single cells of the same cell type is surprisingly high. The authors explore potential scenarios where TSSs observed in single cells are not derived from artifacts but potentially from coordinated bursting. However, additional supporting evidence to demonstrate the co-existence as well as regulatory connections of alternative TSSs in single cells would greatly strengthen the claims made in this manuscript.

Comments

Alternative methods to validate multiple TSSs based on 5'-RACE (e.g. single cell PCR based) or single-molecule FISH of alternative TSSs would strongly substantiate the claim that multiple TSSs co-exists in a single cell. For instance, different fluorescently labeled probes targeting alternative TSSs with larger spans (e.g. ~5 kb) would clearly demonstrate that two transcripts with different TSS co-exist in a single cell.

It would be imperative that the observations made in this study is not affected by potential doublets innate to the microfluidics system used in the original Zeisel a., et al. Science 2015 paper. Is the co-existence of alternative TSSs greater than the expected number of doublets found in the study and can the authors demonstrate statistical significance of its enrichment to single (stand alone) TSSs?

Distinct binding sites of transcription factors often control alternative TSSs. Do authors identify putative transcription factor binding sites (TFBS) in the proximity of alternative TSS that would suggest any possible explanation for its co-activation?

As related to the previous point, figure 4A illustrates that the significant alternative TSS usage is contextual to the cell type. This seems to suggest that differential regulatory elements (TFBS, transcription factors) are regulating cell-specific TSS expression. Do authors find additional evidence of regulatory components in these alternative TSSs (in the same single cells and across different cell types)?

Is there any functional distinction between alternative TSS vs alternative promoters defined in this study?

Reviewer #2:

The authors leverage their recent 5' single cell data of single mouse cortex and hippocampal cells to examine regulation of 5' TSS usage at single cell resolution. They find that the ratio of alternative TSS usage is largely conserved across single cells, and (with a small number of exceptions), found little evidence for stochastic or mutually exclusive TSS selection.

The paper addresses an important and previously unanswered question - the coordination of 5' TSS across single cells within the same type, as well as across cell types. Indeed, the dataset leveraged here is unique in its capability to answer this question, as the STRT technique captures 5' TSS sites, and the use of UMIs enables the author to control for extensive PCR jackpotting that can be highly confounding. However, I was unable to discern a clear message from the manuscript, and have significant concerns about the analytical framework used to assess TSS regulation.

Main comments

1. Overall, I thought the authors did a good job with annotating single cell reads to individual TSS.

This type of analysis has many caveats, but the QC and description shows that this was done carefully and accurately.

2. I did struggle to come away with a clear set of messages from the manuscript. Many of the results are presented anecdotally (i.e.. : Interestingly,.....which may reflect...) and are not followed up on in detail. Examples include:
- The potential for 5' TSS reads to more accurately reflect new transcription
- Enrichment of reads prior to the stop codon, reflecting stalled ribosome dynamics in RNA-seq (?)
- Changes in correlation strength depending on the location of the TSS
- The Cst3 result (more below)

Any one of these anecdotes are potentially interesting with further follow-up analysis, although some seem quite unlikely. As is, they detract from the main story, and make it difficult to focus on the main findings.

It appears that there is a single main finding, that most single cells express multiple TSS and at similar ratios. Its not clear to me that this is particularly unexpected or exciting, but it does contribute to an open question, particularly with recent reports of greater stochasticity in splicing and 3' processing.

3. I disagree strongly with the use of correlation as a metric for measuring TSS usage. The strength of the correlation is interpreted as a measure of the regulation strength of the TSS, but I don't believe that this is a correct interpretation. Instead, the correlation is driven by different overall expression levels of the gene across single cells, perhaps due to differences in cell size or capture efficiency.

For example - suppose as a thought experiment that the authors sequenced 100 cells of identical size . Lets say for gene X total counts were ~50 for each cell. Now suppose that there are two TSS which are nearly 50/50 across each single cells (so each cell detects ~ 28-32 counts of TSS1 and TSS2). If the authors do a correlation plot of TSS1 vs TSS2 here, they will see no correlation, despite, the fact that the TSS choice is incredibly tightly regulated. Therefore, I feel that the correlation-based anlayses are inappropriate to use here.

Instead, the authors should focus on the ratio of TSS usage. They use this analysis sparingly for single cells, particularly later in the manuscript, but the correlation analyses are more prominent. The challenge of using ratios, of course, is that they are difficult for cells with low molecular counts for gene X.

I want to suggest a different analysis that the authors could easily employ, that would significantly improve the manuscript. The authors suggest that a null hypothesis - where each cell expresses the TSS at the same ratio - holds for most genes and most cells. Why not test this hypothesis directly with a binomial (or multinomial) test? The authors could define a TSS ratio (based on pooled data), and then search for cells whose read data deviated from this average ratio. In this case, cells with low counts would not have significant evidence to deviate from the null model, but cells with high molecule counts could significantly deviate if they had skewed read counts. Then, the authors could explicitly test the hypothesis they state, rather than examining it in an indirect and potentially misleading way through correlation analysis.

My intuition is that the conclusions will not change significantly, but I feel this is an important analysis.

3. The Cst3 result is quite interesting, observing a subpopulation of cells that predominantly choose a single TSS. However, the section feels perfunctory, with only two 'variable TSS' genes highlighted, and is missing needed follow-up analyses.
- Are there other genes like CST3 (the binomial test discussed above could easily identify these)
- For the subpopulation that primarily expresses one TSS, is there anything else that is different (either for expression or TSS usage) across single cells between cells?
- is this relationship true across all cell types, or just the two discussed?

4. The comparison of population pools did not significantly add to the manuscript. The main finding seems to be that TSS are used similarly for similar cell types, but there are differences when

considering different tissue cell types (i.e. neuronal and immune cell types). This does not seem surprising or unexpected, and could easily have been observed with bulk data (even with low-medium frequency sorting errors)

1st Revision - authors' response 02 March 2017

Reviewer #1:

The manuscript entitled "Alternative TSSs are co-regulated in single cells in the mouse brain" is a follow up analysis of previously published Zeisel A., et al "Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq". Understanding of TSS usage in single cells and also across different cell types is essential step to dissect the regulatory process surrounding promoters and this manuscript uses the single cell sequencing data to address this important question. The authors report that alternative TSS usage is a regulated process and the correlation between two TSS expressed in single cells of the same cell type is surprisingly high. The authors explore potential scenarios where TSSs observed in single cells are not derived from artifacts but potentially from coordinated bursting. However, additional supporting evidence to demonstrate the co-existence as well as regulatory connections of alternative TSSs in single cells would greatly strengthen the claims made in this manuscript.

Comments:

1.
Alternative methods to validate multiple TSSs based on 5'-RACE (e.g. single cell PCR based) or single-molecule FISH of alternative TSSs would strongly substantiate the claim that multiple TSSs co-exists in a single cell. For instance, different fluorescently labeled probes targeting alternative TSSs with larger spans (e.g. ~5 kb) would clearly demonstrate that two transcripts with different TSS co-exist in a single cell.

Response:
We agree that independent validation would strengthen the findings. Single-molecule FISH would be very challenging since in most cases, alternative TSSs create nested transcripts where the shorter transcript is entirely included in the longer transcript. This makes it difficult to directly detect the shorter transcript by smFISH, independent of the longer transcript (i.e. there is no unique sequence to target). 5'-RACE is essentially equivalent to the single-cell mRNA seq method we used (indeed, Clontech kits for 5'-RACE are based on exactly the same template switching technology). Thus it would not serve as a truly independent validation.

Instead, we opted to perform PacBio full-length sequencing of six single cells. PacBio sequencing confirmed alternative TSS usage in single cells for multiple genes and these findings are now included in supplementary figure 15A-C.

2. It would be imperative that the observations made in this study is not affected by potential doublets innate to the microfluidics system used in the original Zeisel a., et al. Science 2015 paper. Is the co-existence of alternative TSSs greater than the expected number of doublets found in the study and can the authors demonstrate statistical significance of its enrichment to single (stand alone) TSSs?

Answer:
This is an important point, and we have now directly tested the hypothesis that observed alternative TSSs were artefacts due to doublets. We reasoned that if alternative TSSs were predominantly observed in doublets, then genes with cell type-specific expression should show much lower rates of alternative TSS usage than genes with broader expression (given that cell type-specific genes would not be affected by doublets). However, looking only at neuron specific genes (based on a rank-sum test between CA1-neurons and oligodendrocytes), with a total gene expression greater than 1 molecule per cell in CA1 neurons, 1.93% of these genes contained an alternative TSS in CA1 neurons (where both TSS expressed at least 0.5 molecules per cell in CA1 neurons), as compared with all genes, where 1.90% of all genes contained

aTSS. Thus, we see little evidence that doublets creates a large number of false alternative TSS.

Furthermore, all cells included in this study (and our previous work) have been individually imaged at high magnification, and visible doublets have been removed. We estimate that this procedure leads to a residual rate of doublets less than 10%.

3. Distinct binding sites of transcription factors often control alternative TSSs. Do authors identify putative transcription factor binding sites (TFBS) in the proximity of alternative TSS that would suggest any possible explanation for its co-activation?

Answer:
This is an interestng suggestion, and indeed stable differences in the activity of adjacent TSSs would very likely require differences in the local regulatory machinery. However, the vast majority of TSS were located too close to each other to be able to discern if they are regulated by the same or different TF. We have now examined this possibility for the small number of expressed TSSs located far apart (>1 kb) and that showed alternative regulation, but the result was inconclusive. We note that the main finding of the paper is that simultaneously expressed, alternative TSSs, do not show independent regulation and instead maintain fixed ratios. In such cases we would not expect to find distinct TFs regulating each TSS.

However, we have expanded our analysis of the strikingly bimodal regulation of *Cst3*, and found a significant correlation with a number of genes, including the immediate-early transcription factor Fos, which also has a target site near the *Cst3* promoter.

4. As related to the previous point, figure 4A illustrates that the significant alternative TSS usage is contextual to the cell type. This seems to suggest that differential regulatory elements (TFBS, transcription factors) are regulating cell-specific TSS expression. Do authors find additional evidence of regulatory components in these alternative TSSs (in the same single cells and across different cell types)?

Answer:
Fig 4A shows that only a small fraction of genes carry pairs of TSSs that are expressed at different ratios in distinct cell types. Unfortunately, the short distance between TSSs and the low number of positive cases precluded an analysis of local regulatory elements.

As suggested by another referee, we have removed this section, which contains results that could have also been obtained by bulk RNA-seq.

5. Is there any functional distinction between alternative TSS vs alternative promoters defined in this study?

Answer:
We did not find any sharp distinction between two classes of alternative TSSs, but a gradual shift as a function of the inter-TSS distance (e.g. Fig 1d). We therefore used an operational distinction between alternative TSS and alternative promoters, where if the distance was <1kb they were defined as TSS and if the distance was >1kb they were defined as promoters.

Reviewer #2:

The authors leverage their recent 5' single cell data of single mouse cortex and hippocampal cells to examine regulation of 5' TSS usage at single cell resolution. They find that the ratio of alternative TSS usage is largely conserved across single cells, and (with a small number of exceptions), found little evidence for stochastic or mutually exclusive TSS selection.

The paper addresses an important and previously unanswered question - the coordination of 5' TSS across single cells within the same type, as well as across cell types. Indeed, the dataset leveraged here is unique in its capability to answer this question, as the STRT technique captures 5' TSS sites, and the use of UMIs enables the author to control for extensive PCR jackpotting that can be highly

confounding. However, I was unable to discern a clear message from the manuscript, and have significant concerns about the analytical framework used to assess TSS regulation.

Main comments

1. Overall, I thought the authors did a good job with annotating single cell reads to individual TSS. This type of analysis has many caveats, but the QC and description shows that this was done carefully and accurately.

2. I did struggle to come away with a clear set of messages from the manuscript. Many of the results are presented anecdotally (i.e.. : Interestingly,.....which may reflect...) and are not followed up on in detail. Examples include:
- The potential for 5' TSS reads to more accurately reflect new transcription
- Enrichment of reads prior to the stop codon, reflecting stalled ribosome dynamics in RNA-seq (?)
- Changes in correlation strength depending on the location of the TSS
- The Cst3 result (more below)

Any one of these anecdotes are potentially interesting with further follow-up analysis, although some seem quite unlikely. As is, they detract from the main story, and make it difficult to focus on the main findings.

Answer:
We have revised the manuscript to more clearly focus on a few key findings, especially Cst3, and hope that the discussion is now easier to follow.

It appears that there is a single main finding, that most single cells express multiple TSS and at similar ratios. Its not clear to me that this is particularly unexpected or exciting, but it does contribute to an open question, particularly with recent reports of greater stochasticity in splicing and 3' processing.

Answer:
Considering the recent discovery of transcriptional bursting and as reviewer #2 mention the great stochasticity in splicing and 3' processing we think the finding of a fixed ratio between the major and minor TSS to be both rather surprising and of considerable value.

First of all, it informs current work on the 3D architecture of chromatin, and shows that there may be no need to assume that distal enhancers are highly selective for specific individual TSSs or promoters.

Second, it shows that whatever mechanism is responsible for stochastic gene expression cannot be intrinsic to individual TSSs, but must operate at the level of a set of TSSs (at minimum, pairs of TSSs, which is what we studied here).

3. I disagree strongly with the use of correlation as a metric for measuring TSS usage. The strength of the correlation is interpreted as a measure of the regulation strength of the TSS, but I don't believe that this is a correct interpretation. Instead, the correlation is driven by different overall expression levels of the gene across single cells, perhaps due to differences in cell size or capture efficiency.

For example - suppose as a thought experiment that the authors sequenced 100 cells of identical size . Lets say for gene X total counts were ~50 for each cell. Now suppose that there are two TSS which are nearly 50/50 across each single cells (so each cell detects ~ 28-32 counts of TSS1 and TSS2). If the authors do a correlation plot of TSS1 vs TSS2 here, they will see no correlation, despite, the fact that the TSS choice is incredibly tightly regulated. Therefore, I feel that the correlation-based anlayses are inappropriate to use here.

Instead, the authors should focus on the ratio of TSS usage. They use this analysis sparingly for single cells, particularly later in the manuscript, but the correlation analyses are more prominent. The challenge of using ratios, of course, is that they are difficult for cells with low molecular counts for gene X.

I want to suggest a different analysis that the authors could easily employ, that would significantly improve the manuscript. The authors suggest that a null hypothesis - where each cell expresses the TSS at the same ratio - holds for most genes and most cells. Why not test this hypothesis directly with a binomial (or multinomial) test? The authors could define a TSS ratio (based on pooled data), and then search for cells whose read data deviated from this average ratio. In this case, cells with low counts would not have significant evidence to deviate from the null model, but cells with high molecule counts could significantly deviate if they had skewed read counts. Then, the authors could explicitly test the hypothesis they state, rather than examining it in an indirect and potentially misleading way through correlation analysis.

My intuition is that the conclusions will not change significantly, but I feel this is an important analysis.

Answer:
We thank you for this valuable suggestion. This analysis has now been carried out and your intuition was correct, the conclusion didn't change but was rather confirmed. Cst3 still stands out as a rare exception to the hypothesis that the ratio between the major and minor TSS is constant. A figure has been added to the main text (figure 2F) as well as supplemental figure 12. The text describing the binomial test is on page 8, paragraph 2 and well as in the methods section.

3. The Cst3 result is quite interesting, observing a subpopulation of cells that predominantly choose a single TSS. However, the section feels perfunctory, with only two 'variable TSS' genes highlighted, and is missing needed follow-up analyses.
- Are there other genes like CST3 (the binomial test discussed above could easily identify these)
- For the subpopulation that primarily expresses one TSS, is there anything else that is different (either for expression or TSS usage) across single cells between cells?
- is this relationship true across all cell types, or just the two discussed?

Answer:
Further investigations into the bimodal expression of Cst3 have been carried out. For most cell types Cst3 was the only gene with bimodal expression, however for Oligodendrocytes there are a few more genes with a prominent number of cells deviating from the expected major to minor TSS ratio, most notably Plp1 (see supporting figure 12)

Genes with differential expression between cells with high and low Cst3 major TSS expression have been identified as well as GO terms associated with those genes (supplemental table 3). It seems that cells with high Cst3 major TSS expression are more involved in inter-cellular communication. Significant GO terms include receptor activity, extracellular space and plasma membrane. A supporting table with GO terms defining high and low Cst3 major TSS expression for the four cell types with bimodal Cst3 TSS expression have been added (supporting table 3)

4. The comparison of population pools did not significantly add to the manuscript. The main finding seems to be that TSS are used similarly for similar cell types, but there are differences when considering different tissue cell types (i.e. neuronal and immune cell types). This does not seem surprising or unexpected, and could easily have been observed with bulk data (even with low-medium frequency sorting errors)

Answer:
We agree that these results could have in principle been obtained from bulk data. We have removed this section, making room for the more extensive analysis of *Cst3* on page 9-10.

2nd Editorial Decision 03 April 2017

Thank you again for sending us your revised manuscript. We have now heard back from the two referees who were asked to evaluate the study. As you will see below, both referees think that the previously raised issues have been satisfactorily addressed and the study is now suitable for publication.

Before we formally accept the manuscript we would ask you to address some remaining editorial issues listed below.

REFEREE REPORTS

Reviewer #1:

The description of alternative TSS in single cells is well presented and opens new avenues to explore gene regulation at the promoter level. The authors have adequately addressed comments raised by this reviewer.

Reviewer #2:

The authors have addressed my comments with strong revisions that have substantially improved the manuscript.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name:	
Journal Submitted to:	
Manuscript Number:	

Reporting Checklist For Life Sciences Articles (Rev. July 2015)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

Please ensure that the answers to the following questions are reported in the manuscript itself. We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

In the pink boxes below, provide the page number(s) of the manuscript draft or figure legend(s) where the information can be located. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable).

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	This is an explorative study of preexisting previously published data. We used the most comprehensive single-cell RNA 5' sequencing data-set available at the time.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	Not used
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	Seven cell types with reasonable cell numbers were included
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	We used no treatment condition
For animal studies, include a statement about randomization even if no randomization was used.	Not used
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	Group allocation was taken from previously published data
4.b. For animal studies, include a statement about blinding even if no blinding was done	Not used
5. For every figure, are statistical tests justified as appropriate?	Yes
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Yes. If sample size was small (n<20) a non-parametric test was used. Also if median better represents the distribution of the data than the mean a non-parametric test was used. The method to assess this was to plot the distributions.
Is there an estimate of variation within each group of data?	Yes
Is the variance similar between the groups that are being statistically compared?	Yes, except in one case where, but there combined sample sizes are large (n>70)

C- Reagents

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	Not used
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	Not used

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	The single cell RNA-seq data came from a previously publihsed data set using genetically outbred (CD-1) mice.
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	We used previously published data
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	We used previously published data

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	No human subject was used
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	

F- Data Accessibility

18. Provide accession codes for deposited data. See author guidelines, under 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	The data is accessible at the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) under accession code GSE60361
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	The data is accessible at the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) under accession code GSE60361
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	No human data used
21. As far as possible, primary and referenced data should be formally cited in a Data Availability section. Please state whether you have included this section. Examples: **Primary Data** Wetmore KM, Deutschbauer AM, Price MN, Arkin AP (2012). Comparison of gene expression and mutant fitness in Shewanella oneidensis MR-1. Gene Expression Omnibus GSE39462 **Referenced Data** Huang J, Brown AF, Lei M (2012). Crystal structure of the TRBD domain of TERT and the CR4/5 of TR. Protein Data Bank 4O26 AP-MS analysis of human histone deacetylase interactions in CEM-T cells (2013). PRIDE PXD000208	This section is included
22. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	No computational model created

G- Dual use research of concern

23. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	No risk of dual use research of concern